BEST PARAMETER CHOICE OF STOCHASTIC RESONANCE TO ENHANCE FAULT SIGNATURE IN BEARINGS

Marchesiello, S., Fasana, A., Garibaldi, L.

Dipartimento di Ingegneria Meccanica e Aerospaziale
Politecnico di Torino
Corso Duca degli Abruzzi 24, 10129, Torino, Italy
stefano.marchesiello@polito.it

SUMMARY: Stochastic Resonance (SR) is a phenomenon studied and exploited for telecommunication, which permits the detection and amplification of weak signals by the assistance of noise. The first papers on this topic date back to the early 80s and were developed to explain some periodic natural phenomena. Other applications are in neuroscience, biology, medicine and, obviously, mechanics.
Recently, a few researchers have tried to apply this technique for detecting faults in mechanical systems and also bearings. In this paper we discuss the best way to select the parameters to augment the performance of the algorithm. This is probably the main drawback of SR, since in system identification the procedure should be as blind as possible to be efficient and widely applicable. The classical bi-stable potential form is adopted in our study, with three parameters to be selected. Based on numerical tests, a characteristic trend of the amplification factor has been found with respect to the parameters variation, so that a general rule is consequently determined which gives the best performances in terms of detection and amplification. The SR algorithm is tested on both simulated and experimental data showing a good capacity of increasing the signal to noise ratio.

KEYWORDS: stochastic resonance, bearing diagnostics, damage detection.

1. INTRODUCTION

The phenomenon of stochastic resonance (SR) has applications in a number of different fields and scientific domains. The possibility of resonance in dynamical systems, which behave stochastically, was introduced by R. Benzi et al. [1] in 1981 and originally exploited for studying the evolution of the Earth's climate. Its first applications were in a wide range of problems connected to physical and life sciences. Other observations of this phenomenon concern experiments on electronic circuits, chemical reactions, semiconductor devices, nonlinear optical systems, magnetic systems and superconducting quantum interference devices [2].
The studies of SR for mechanical applications, especially for mechanical fault diagnosis, began in the mid-90s and great improvements have been achieved in particular during the last years. Several techniques exist and are applied for the detection of defects in rotating machines such as gears or bearings in many industrial applications, but SR is the only one that takes advantage of noise. In fact mechanical acquisitions are generally strongly corrupted by background noise from other elements of the system: this component is usually neglected, but on the contrary it is used by SR to enhance the features of faults [3].

2. STOCHASTIC RESONANCE

Stochastic Resonance is a tool used in signal processing to increase the signal-to-noise ratio (SNR) of the output of a non-linear dynamic system, in order to extract the characteristic features of the system under investigation

from background noise. This is obtained by adding a nonlinear dynamic system to the measured signal corrupted by noise and, by properly tuning, enhancing the signal of interest (Figure 1).
Usually noise is considered as a disturbance that may make measured data unusable, so that the basic idea behind each data processing procedure includes the filtering or removal of noise. However, useful information may happen to be corrupted or destroyed by this procedure, so much attention has to be paid. In SR, instead, noise is a basic element of the process: in fact, according to the classical theory of SR, by adding a well-tuned noise to the full measurement, signal detection is facilitated.
The amplification of weak signals is obtained by varying the noise level, through the addition of a potential, but keeping the input modulation signal. SR mechanism implies that, if a sinusoidal driving frequency mixed with noise is given as input to a nonlinear system, its output contains a high peak in the spectrum corresponding to the driving frequency which varies its amplitude as a function of the noise added through the system [3-5].
The algorithms for SR, especially for weak impulses or aperiodic signals, work in time domain and several types of implementation exist through the use of different kind of potentials.

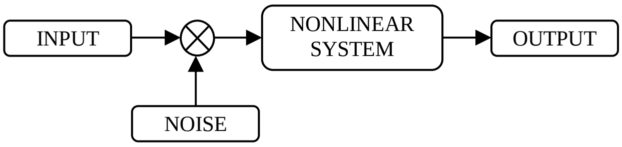


Figure 1 – Stochastic resonance scheme

The dynamic behavior of SR can be described by the following Brownian motion equation of particles, where $s(t)$ and $n(t)$ are respectively the input signal and the noise, whose sum is in practice the measurement signal, and $V(x)$ is the potential function:

$$\frac{dx}{dt} = -\frac{dV}{dx} + s(t) + n(t) \tag{1}$$

The quantity $x(t)$ is the system output and denotes the trajectory of the Brownian particle in the potential function $V(x)$. Classical SR [3-4] uses the following polynomial expression as a potential function:

$$V(x) = -\frac{a}{2}x^2 + \frac{b}{4}x^4 \tag{2}$$

It represents a bi-stable symmetric system where a and b are positive real parameters, whose two stable points are located at $x_m = \pm\sqrt{a/b}$ and the height of the potential barrier is $\delta V = a^2/4b$. Substituting this in the Brownian particle motion equation and considering a periodic signal $s(t)$ of amplitude A and driving frequency f_0 and a Gaussian white noise $n(t)$ with zero mean and given variance, the main equation of the process is obtained:

$$\frac{dx}{dt} = ax - bx^3 + A_0\cos(2\pi f_0 t + \varphi) + n(t) \tag{3}$$

In case there is no external excitation the position is only determined by the initial conditions and never changes. If a periodic input function at frequency f_0 is given as input, the potential function is modulated and changes periodically, and in case there is also noise in the input, the particle will jump between the potential wells with a period corresponding to that of the input function (Figure 2). So, by properly tuning the potential to the noise present in the signal it is possible to detect weak signals by simply solving the above first-order differential equation using the discrete Runge–Kutta method.
At this stage it is necessary to find the correct values of a and b potential parameters so that $x(t)$ takes the form of a wave with the same oscillation frequency as the driving frequency of the periodic signal. In fact, the system output $x(t)$, which represents the motion of a Brownian particle inside the potential $V(x)$, should oscillate between the two potential wells at a transition rate that matches the period of the input signal. Consequently the periodic input signal is enhanced only by adjusting the dynamic system parameters.
In a more realistic case of impact signals given as input s(t) to the system, the output x(t) will be made of a series of impulses located in the exact position as the original signal. For example, by considering only a single impact event, the equation of motion becomes:

$$\frac{dx}{dt} = ax - bx^3 + Ae^{-Dt}\sin(2\pi f_0 t) + n(t) \tag{4}$$

In this case the Brownian particle can jump between the potential wells in a few oscillation periods or just in one by properly tuning the potential parameters, while in the remaining parts of the signal it will get stuck inside the potential well because no sufficient energy is provided by noise.

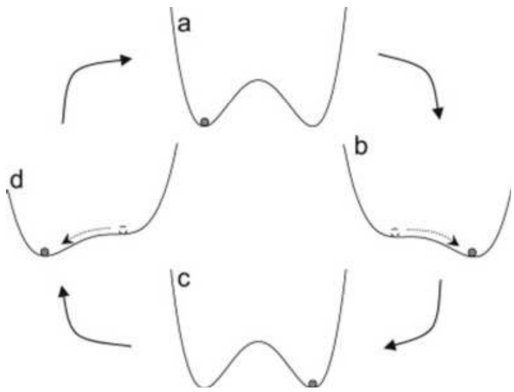


Figure 2 – State transition of the bistable system in the presence of periodic input and noise: (a) initial state when t=0; (b) t =1/4f_0; (c) t=1/2f_0; (d) =3/4f_0. Figure from [4]

2.1. The algorithm for impact signal detection

The main difficulty for an efficient implementation of SR is the selection of the parameters of the potential. In fact, a criterion is necessary to determine if the selected set of values gives good results in the output, which means its capability of enhancing signal.

Several measurement indexes exist to assess the quality of the procedure, for example kurtosis, crest factor or others. The first is defined as the ratio between the fourth central moment and the square of the variance and it is a measure of the peakness of a probability distribution of a real-valued random variable. The more the peaks are narrow and sharp the more kurtosis is high, in contrast to the case in which there is a normal distribution when the kurtosis tends to 3:

$$kurt(x) = \frac{E[(x-\bar{x})^4]}{(E[(x-\bar{x})^2])^2} \tag{5}$$

Crest factor is defined as the absolute peak value over the root mean square of the distribution. The higher are the peaks emerging from the background noise after the application of the method, the more this factor increases.

$$CF = \frac{|x|_{peak}}{x_{RMS}} = \frac{|x|_{peak}}{\sqrt{E[x^2]}} \tag{6}$$

Whatever index is selected, the parameters of the potential have to be selected in order to maximize it.

The first step after signal pre-processing is the initialization of the range of the coefficients in terms of minimum and maximum values but also step size. By substituting them in the SR equation, the output signal is computed with a fourth-order Runge–Kutta algorithm and then evaluated through the criterion previously selected. Generally, after the solution of each equation is obtained, the transient response is removed in order to correctly compute the corresponding criterion function, otherwise numerical peaks could be included in its evaluation. Once all possible values of the coefficients have been used, the maximum value of the criterion is chosen together with the corresponding values of the coefficients. The improved waveform is then computed and, from this, it is possible to get all the information on the characterization of the impacts.

2.2. Re-scaling ratio

The main problem of the classical bi-stable stochastic resonance is the small parameters restriction, which is in contrast with bearing fault diagnostics. Essentially SR focuses on low frequency and weak periodic signal submerged in small noise, this meaning that the values of the frequency and amplitude of periodic signal and

noise intensity are less than 1. But defect frequencies are usually much higher than 1 Hz and as a consequence it is necessary to adapt SR algorithm to large parameters signals.
The approach usually adopted makes use of a re-scaling factor applied to the sampling frequency f_S in order to make it much lower by linearly compressing the frequency scale. R is a rescaling ratio that satisfies the requirement of small parameters [6].

$$f_{Sr} = f_S/R \tag{7}$$

3. NUMERICAL SIMULATIONS

The SR method has been tested on several numerical data. A first simulation has been performed to show the benefits of the SR approach on a dataset generated by summing an impulse response function (IRF) to a Gaussian noise:

$$s(t) + n(t) = Ae^{-D(t-T_i)} \sin\big(2\pi f_0(t - T_i)\big) + \mathcal{N}(0, \sigma) \tag{8}$$

Figure 3 shows the output produced by SR with the following parameters: natural frequency $f_0 = 16\,Hz$, attenuation rate $D = 12\,s^{-1}$, impulse starting time T_i = 1 s and amplitude A = 0.15. A Gaussian white noise $n(t)$ with zero mean value and standard deviation σ = 0.07 has been added to the IRF. The sampling frequency is set to $f_S = 500\,Hz$, N = 1000 is the number of considered samples and the rescaling ratio is R = 200. Fig. 3 shows that for a particular realization of the noise and by selecting a = 0.7 and b = 9 the SR process can increase the kurtosis index from 3 to 29 and the presence of an IRF is clearly detectable.

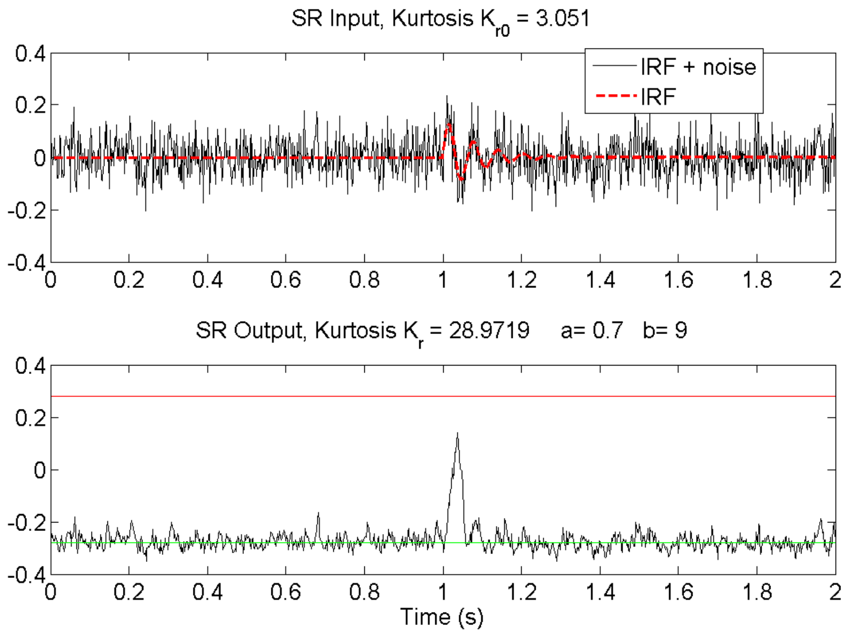


Figure 3 – Example of SR effect: the Kurtosis of the output signal has been considerably increased.

It also results that if the height of the potential barrier is too low the Brownian particle can jump between the potential wells many times, and not only when the actual impact occurs. In this case, the capability of identifying the actual IRF is lost and the Kurtosis index of the SR output doesn't show any increase. To demonstrate this phenomenon, 500 realizations of the Gaussian noise have been considered with the above-mentioned parameters. For each realization, parameter *a* varies between 0.1 and 1 and parameter *b* between 1 and 10. The corresponding variation steps are defined in such a way to obtain a 10×10 search grid. Results of maximum, minimum and mean values of the Kurtosis index computed over these realizations are depicted in Fig. 4a as a function of the height of the potential barrier $\delta V = a^2/4b$: this figure shows that an optimum value of δV exists.
To study the IRF amplitude *A* influence on the Kurtosis gain, the Kurtosis index is depicted in Figure 4b with several amplitudes *A*: this figure confirms that an optimum value of the height of potential barrier δV exists and that this value moves to the right if the IRF amplitude grows. It is stressed that this amplitude is related to the damage severity.

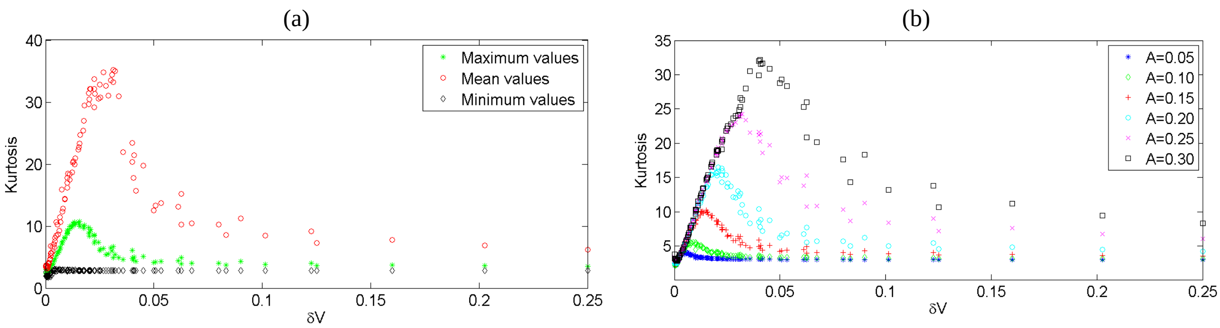
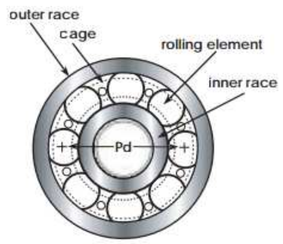


Figure 4 –Kurtosis of the output signal as a function of the height of the potential barrier (500 simulations and 10×10 search grid): (a) A = 0.15, (b) A from 0.05 to 0.30

4. SIMULATIONS OF BEARING FAULTS

Bearings can manifest several kinds of damage due to different causes, such as fatigue, wear, poor installation, improper lubrication and occasionally manufacturing faults. Defects could arise in all constituting elements and each has a distinct pattern in the time signal acquisition and could be identified by its deep examination. Bearing is made of the following components: outer race, inner race, cage and rolling elements (Table 1).

Table 1 – Bearing elements and defect frequencies.

	Ball Pass Frequency of the Outer race	$BPFO = \frac{nf_r}{2}\left(1-\frac{d}{D}cos\phi\right)$
	Ball Pass Frequency of the Inner race	$BPFI = \frac{nf_r}{2}\left(1+\frac{d}{D}cos\phi\right)$
	Fundamental Train Frequency (cage speed)	$FTF = \frac{f_r}{2}\left(1-\frac{d}{D}cos\phi\right)$
	Ball Spin Frequency	$BSF = f_r\frac{D}{2d}\left(1-\left(\frac{d}{D}cos\phi\right)^2\right)$

When a bearing spins, any irregularity in the surface of inner or outer race, or in the roundness of the rolling elements excites periodic frequencies called fundamental defect frequencies. These depend on the geometry of the bearing and clearly on the shaft speed. In Table 1 there is a list of these frequencies in which d is ball diameter, D is pitch diameter, ϕ is contact angle, f_r is shaft speed, n is the number of rolling elements. It is assumed that outer race is fixed and inner race rotates.

All previously listed formulas are valid only when pure rolling contact exists between balls, inner race and outer race, but actually there is always some random slip when a bearing is under load and after some wear and consequently frequencies are not found exactly as predicted by Table 1.

In this paper faulty bearing time histories are generated through a Matlab code as repeated impulse response functions, submerged by background noise. These data are given in input to the SR algorithm for signal fault enhancement. The simulated bearing has the characteristics listed in Table 2, where its defect frequencies are computed as function of shaft speed.

The standard deviation of noise was set to σ = 0.07 and the defect was simulated as the impulse response of a SDOF system with resonance frequency of 5500 Hz and damping ratio 5%. In the simulation, which is carried out at f_S = 96000 Hz for 1 s, the shaft speed is 100 Hz, and the amplitude of the IRF is set at A = 0.30. The model also includes the typical modulation for unidirectional load, which is at the cage speed for rolling element faults. A re-scaling factor R=20000 is used in order to satisfy the small parameter requirement. By varying the a and b parameters in the same grid as in the previous example, it was found that the dependence of the Kurtosis index on the height of the potential barrier shows a "resonance curve" which is similar to Fig. 4a. The optimal value of δV produces a Kurtosis gain of about 6 as shown in Fig. 5a. To demonstrate the effectiveness of the SR algorithm the normalized power density spectrum of both input and output is depicted in Fig. 5b: the defect frequency peak is increased by one order of magnitude. Note that the ballspin frequency (BSF) is the frequency of fault passage over the same race (inner or outer), so that in general there are two shocks per basic period.

Table 2 – Bearing dimensions and defect frequencies as functions of the shaft frequency.

d	D	ϕ	n	$BPFO$	$BPFI$	FTF	BSF
$9\ mm$	$40.5\ mm$	$0°$	10	$3.89\ f_r$	$6.11\ f_r$	$0.39\ f_r$	$2.14\ f_r$

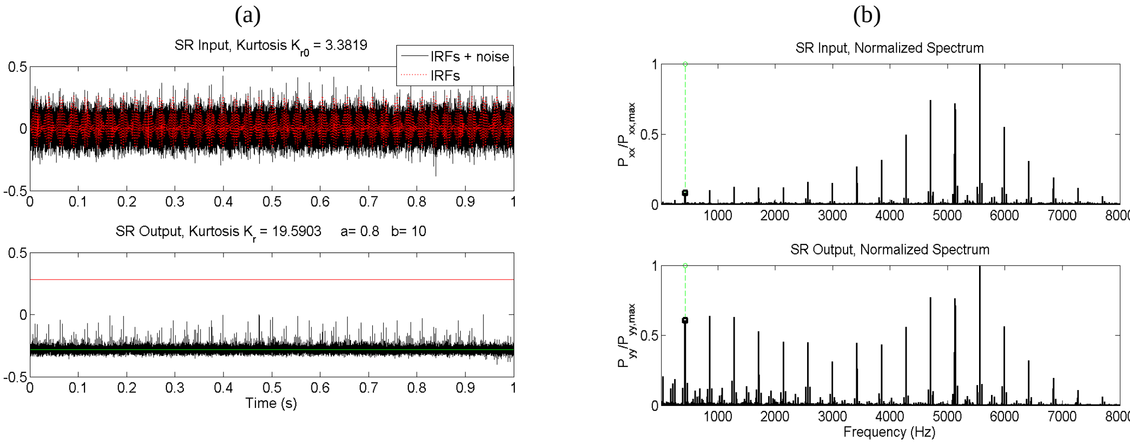


Figure 5 – Example of SR effect (bearing model): (a) the Kurtosis of the output signal and (b) the peak corresponding to the defect frequency have been considerably increased.

5. EXPERIMENTAL CASE

Finally an experimental application of SR is presented in order to show its effectiveness. The test rig, set up in the laboratory of the Department of Mechanical and Aerospace Engineering of Politecnico di Torino, is made of three bearings and a rotating shaft, see Figure 6. The radial load is applied to the central bearing while the other two serve as supports for the shaft: one of the lateral bearing exhibits different levels of damage. Damage monitoring is performed by equipping the structure with triaxial accelerometers at different shaft speeds, load levels and oil temperature.

The bearing under exam has the characteristics and defect frequencies as in Table. 2. Results shown hereafter concern a bearing with a defect in a rolling element, with shaft speed of $278\ Hz$. The central support is loaded by a force of $1800\ N$ and the sampling frequency is $f_S = 51200\ Hz$.

The optimal value of δV produces a Kurtosis gain of about 7 as shown in Fig. 7a. The normalized power density spectrum of both input and output is depicted in Fig. 7b: the defect frequency peak was not detectable in the raw spectrum, while it clearly appears in the SR output together with its harmonics. The identified defect frequency differs from the theoretical value (Tab. 2) by an amount of 3%, which is a typical value of change in bearing defect frequencies.

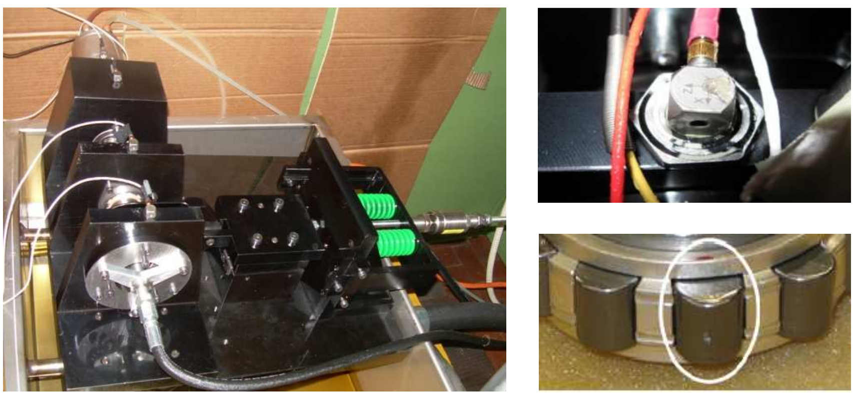


Figure 6 – The test rig, the axes orientation of the triaxial accelerometers and the damaged roller used in the tests.

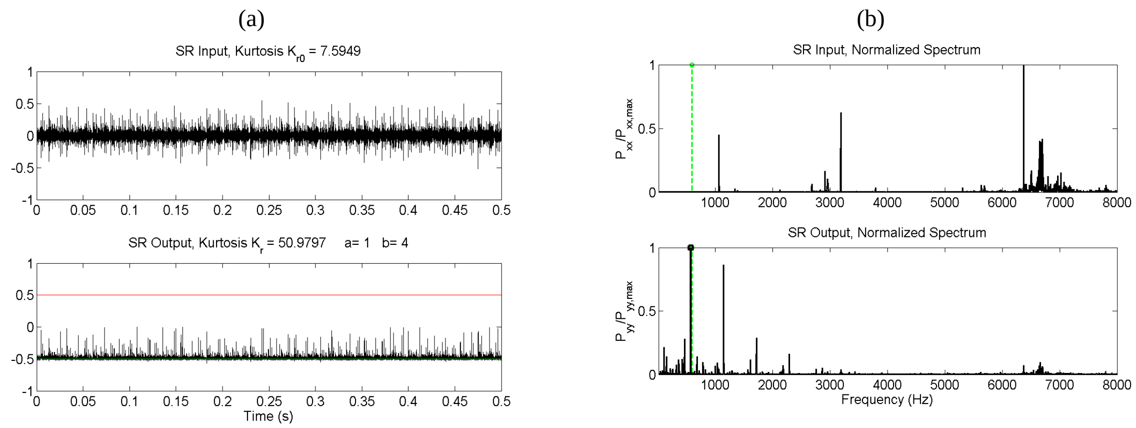


Figure 7 – Example of SR effect (experimental case): (a) the Kurtosis of the output signal has been considerably increased and (b) the peak corresponding to the defect frequency and its harmonics have been highlighted.

6. CONCLUSIONS

The mathematics behind the mechanism of stochastic resonance is relatively simple and easy to be implemented. Theoretically SR works really well in the detection of pulses submerged by background noise even with low levels of excitation. However, the main limitation of the procedure is the choice of the parameters *a* and *b*. In this paper it is shown that an optimum value of the height of the potential barrier δV exists, which produces a considerable increase in the Kurtosis index. This property reduces the user-selected parameters from two (*a* and *b*) to one (δV), thus decreasing dramatically the computational effort in the parameter tuning process. This computational time reduction is of paramount interest in the automatic health monitoring.

AKNOWLEDGEMENTS

This work has been conducted in the framework of the project GREAT 2020 (fase II) – piattaforma Aerospazio.

7. REFERENCES

[1] Benzi, R, Sutera, A., Vulpiani A, “The mechanism of stochastic resonance”, Journal of Physics A: mathematical and general 14, 1981, L453-L457.

[2] McDonnell, MD. , Abbott, D.,”What Is Stochastic Resonance? Definitions, Misconceptions, Debates, and Its Relevance to Biology”, PLoS Computational Biology 5(5), 2009.

[3] Li, J., Chen, X., He, Z., “Multi-stable stochastic resonance and its application research on mechanical fault diagnosis”, Journal of Sound and Vibration 332, 2013, pp. 5999-6015

[4] Li, J., Chen, X., He, Z., “Adaptive stochastic resonance method for impact signal detection based on sliding window”, Mechanical Systems and Signal Processing 36, 2013, pp. 240-255

[5] Lu, S., He, Q., Kong, F., “Stochastic resonance with Woods–Saxon potential for rolling element bearing fault diagnosis”, Mechanical Systems and Signal Processing 45, 2014, pp. 488-503

[6] Tan, J., Chen, X., Wang, J., Chen, H., Cao, H., Zi, Y., He, Z., “Study of frequency-shifted and re-scaling stochastic resonance and its application to fault diagnosis”, Mechanical Systems and Signal Processing 23, 2009, pp. 811-822